INTRODUCTION TO MATRIX ALGEBRA

1. Definition of a matrix and a vector

1.1. **Definition of a matrix.** A matrix is a rectangular array of numbers arranged into rows and columns. It is written as

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1}$$

The above array is called an m by n (m x n) matrix since it has m rows and n columns. Typically upper-case letters are used to denote a matrix and lower case letters with subscripts the elements. The matrix A is also often denoted

$$A = \|a_{ij}\| \tag{2}$$

1.2. **Definition of a vector.** A vector is a n-tuple of numbers. In R^2 a vector would be an ordered pair of numbers $\{x,\, y\}$. In R^3 a vector is a 3-tuple, i.e., $\{x_1,\, x_2,\, x_3\}$. Similarly for R^n. Vectors are usually denoted by lower case letters such as a or b, or more formally $\vec{a}$ or $\vec{b}$.

1.3. **Row and column vectors.**

1.3.1. *Row vector.* A matrix with one row and n columns (1xn) is called a row vector. It is usually written $\vec{x}\,'$ or

$$\vec{x}\,' = \begin{pmatrix} x_1, & x_2, & x_3, \dots, & x_n \end{pmatrix} \tag{3}$$

The use of the prime $'$ symbol indicates we are writing the n-tuple horizontally as if it were the row of a matrix. Note that each row of a matrix is a row vector.

1.3.2. *Column vector.* A matrix with one column and n rows (nx1) is called a column vector. It is written as

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ . \\ x_n \end{pmatrix} \tag{4}$$

Note that each column of a matrix is a column vector. It is common to write the columns of a matrix as $a_1, a_2, \dots a_n$ where each column vector a_j is of length m. As an example a_2 is given by

Date: August 27, 2004.

$$\vec{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ . \\ . \\ . \\ a_{m2} \end{pmatrix} \tag{5}$$

2. Various types of Matrices and Vectors

2.1. **Square matrices.** A square matrix is a matrix with an equal number of rows and columns, i.e. m=n.

2.2. **Transpose of a matrix.** The transpose of a matrix A is a matrix formed from A by interchanging rows and columns such that row i of A becomes column i of the transposed matrix. The transpose is denoted by A' or A^T and

$$A' = \|a_{ji}\| \; when \; A = \|a_{ij}\| \tag{6}$$

If a'_{ij} is the ij$_{th}$ element of A', then $a'_{ij} = a_{ji}$. If the matrix A is given by

$$A = \begin{pmatrix} 3 & 2 & 5 & 7 \\ 1 & 4 & 6 & 3 \\ 5 & 10 & -2 & 0 \\ 1 & 1 & 15 & -2 \end{pmatrix} \tag{7}$$

then A' is given by

$$A' = \begin{pmatrix} 3 & 1 & 5 & 1 \\ 2 & 4 & 10 & 1 \\ 5 & 6 & -2 & 15 \\ 7 & 3 & 0 & -2 \end{pmatrix} \tag{8}$$

2.3. **Symmetric matrix.** A symmetric matrix is a square matrix A for which

$$A = A' \tag{9}$$

An example of a symmetric matrix is

$$T = \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix} \tag{10}$$

$$T' = \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix}$$

2.4. **Identity matrix.** The identity matrix of order n written I or I_n, is a square matrix having ones along the main diagonal (the diagonal running from upper left to lower right and zeroes elsewhere).

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \tag{11}$$

If we write I = $\|\delta_{ij}\|$ then

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{12}$$

The symbol δ_{ij} is called the Kronecker delta. Note that for a system of n equations in n unknowns that has a unique solution, the coefficient matrix of the system after performing the appropriate number of row and column operations is an identity matrix.

2.5. **Scalar matrix.** For any scalar λ, the square matrix

$$S = \| \lambda \, \delta_{ij} \| = \lambda I \tag{13}$$

is called a scalar matrix. An example is

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \tag{14}$$

2.6. **Diagonal matrix.** A square matrix

$$D = \| \lambda_i \, \delta_{ij} \| \tag{15}$$

is called a diagonal matrix. Notice that λ_i varies with i. An example is

$$\begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 56 \end{pmatrix} \tag{16}$$

If a system of equations was written with this coefficient matrix, we could solve the system by solving each equation individually.

2.7. **Null or zero matrix.** The null or zero matrix is a matrix with each element being zero. It is denoted as **0**.

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \tag{17}$$

2.8. **Upper triangular matrix.** A matrix with all elements below the main diagonal equal to zero is called an upper triangular matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ . & & & . & \\ . & & & . & \\ . & & & . & \\ 0 & 0 & 0 & \dots & a_{mn} \end{pmatrix} \tag{18}$$

Specifically $a_{ij} = 0$ if $i > j$ as long as $i < m$ and $j < n$.

2.9. **Lower triangular matrix.** A matrix with all elements above the main diagonal equal to zero is called a lower triangular matrix.

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ . & & & . & \\ . & & & . & \\ . & & & . & \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \tag{19}$$

Specifically $a_{ij} = 0$ if $i < j$ as long as $i < m$ and $j < n$.

3. A NOTE ON SUMMATION NOTATION

3.1. **Single sums.**

3.1.1. *Definition of a single sum.*

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + \dots + a_n \tag{20}$$

3.1.2. *Properties of a single sum.*

$$\sum_{i=1}^{n} ka_i = k \sum_{i=1}^{n} a_i \tag{21}$$

$$\sum_{i=1}^{n} k = k + k + k + \dots + k = nk$$

$$\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$$

3.2. **Double sums.**

3.2.1. *Definition of a double sum.*

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m} a_{1j} + \sum_{j=1}^{m} a_{2j} + \ldots + \sum_{j=1}^{m} a_{nj} \tag{22}$$

$$\begin{aligned}
&= a_{11} + a_{12} + a_{13} + \ldots + a_{1m} \\
&+ a_{21} + a_{22} + a_{23} + \ldots + a_{2m} \\
&+ \ldots \\
&+ \ldots \\
&+ \ldots \\
&+ a_{n1} + a_{n2} + a_{n3} + \ldots + a_{nm}
\end{aligned}$$

3.2.2. *Properties of a double sum.*

$$\begin{aligned}
\Big(\sum_{j=1}^{n} a_j\Big)\Big(\sum_{i=1}^{n} a_i\Big) &= \sum_{i=1}^{n} a_i^2 + 2\sum\sum_{i<j} a_i a_j \\
&= \sum_{i=1}^{n} a_i^2 + \sum\sum_{i\neq j} a_i a_j
\end{aligned} \tag{23}$$

4. Matrix operations

4.1. **Scalar multiplication (matrix).** Given a matrix A and a scalar λ, the product of λ and A, written λA, is defined to be

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix} \tag{24}$$

4.2. **Scalar multiplication (vector).** Given a column vector $\vec{a}$ and a scalar λ, the product of λ and $\vec{a}$, written $\lambda\,\vec{a}$, is defined to be

$$\lambda\,\vec{a} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ . \\ . \\ . \\ \lambda a_m \end{pmatrix} \tag{25}$$

For the second column of a matrix we could write

$$\lambda \vec{a}_2 = \begin{pmatrix} \lambda a_{12} \\ \lambda a_{22} \\ \lambda a_{32} \\ . \\ . \\ . \\ \lambda a_{m2} \end{pmatrix} \tag{26}$$

4.3. **Trace of a square matrix.** The trace of a matrix is the sum of the diagonal elements and is denoted tr A. Consider the matrix C below.

$$C = \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix} \tag{27}$$

The trace of C is [3 + 4 + -2 + -2] = 3.

4.4. **Addition of vectors.** - The sum c of a vector a with m elements and a vector b having m elements is a vector with m elements and whose elements are given by

$$c_j = a_j + b_j \; \forall \, j \tag{28}$$

This gives

$$\vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_m \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_m \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \cdot \\ \cdot \\ \cdot \\ a_m + b_m \end{pmatrix} \tag{29}$$

4.5. **Linear combinations of vectors.** If a and b are two n-vectors and s and t are two real numbers, tz + sb is said to be the linear combination of a and b. In symbols we write,

$$t \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_m \end{pmatrix} + s \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix} = \begin{pmatrix} t\,a_1 + s\,b_1 \\ t\,a_2 + s\,b_2 \\ \cdot \\ \cdot \\ \cdot \\ t\,a_m + s\,b_m \end{pmatrix} \tag{30}$$

Consider three vectors, each with two elements denoted. Call the vectors $\vec{a}_1$, $\vec{a}_2$ and $\vec{b}$. Call the elements of the first one a_{11} and a_{21}, the elements of the second one a_{12} and a_{22} and the elements of $\vec{b}$, b_1 and b_2. Now consider two scalars denoted x_1 and x_2. Now multiply $\vec{a}_1$ by x_1 and $\vec{a}_2$ by x_2 and add the products. We obtain

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}\,x_1 \\ a_{21}\,x_1 \end{pmatrix} + \begin{pmatrix} a_{12}\,x_2 \\ a_{22}\,x_2 \end{pmatrix} = \begin{pmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}x_1 + a_{22}\,x_2 \end{pmatrix} \tag{31}$$

If set this expression equal to $\vec{b}$ we obtain

$$\begin{pmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}x_1 + a_{22}\,x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{32}$$

which is a linear system of 2 equations in 2 unknowns. We can write a general system of m equations in n unknowns as

$$x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + \cdots + x_n\,\vec{a}_n = \vec{b} \tag{33}$$

where x_i are a series of scalar unknowns and each a_j is a column of the A matrix of coefficients.

4.6. **Addition of matrices.** The sum C of a matrix A having m rows and n columns and a matrix B having m rows and n columns is a matrix having m rows and n columns whose elements are given by

$$c_{ij} = a_{ij} + b_{ij}\ \forall\, i, j \tag{34}$$

This gives

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} \quad (35)$$

$$(36)$$

$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ . & & & . \\ . & & & . \\ . & & & . \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix} \quad (37)$$

4.7. **Inner (dot) product of two vectors.** The inner (scalar or dot) product to two vectors u,v of length n is the scalar quantity denoted by

$$u \cdot v = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \dots + u_n v_n \quad (38)$$

4.8. **Multiplication of matrices.** Given an mxn matrix A and an nxr matrix B, the product AB is defined to be an mxr matrix C, whose elements are computed from the elements of A,B according to

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj},\ i = 1, \dots, m,\ j = 1, \dots, r. \quad (39)$$

In other words to obtain the ijth element of c we take the ith row of A and jth column of B and form the inner product. As an example consider the matrices below

$$A = \begin{pmatrix} 3 & 4 & 7 \\ 2 & 5 & 2 \\ 1 & 0 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 4 & 1 \end{pmatrix} \quad (40)$$

The element c_{11} comes from multiplying the first row of A with the first column of B as follows:

$$c_{11} = \begin{pmatrix} 3 & 4 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 3 + 8 + 7 = 18 \quad (41)$$

Similarly the element c_{32} comes from multiplying the third row of A with the second column of B as follows:

$$c_{32} = \begin{pmatrix} 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} = 0 + 0 + 16 = 16 \quad (42)$$

Multiplying out the rest of the entries gives

$$C = \begin{pmatrix} 18 & 32 & 14 \\ 14 & 13 & 9 \\ 5 & 16 & 5 \end{pmatrix} \tag{43}$$

4.9. **Some properties of matrix operations.** Let α and β denote real numbers (scalars), $\vec{a}, \vec{b}, \vec{c}$ denote n-vectors, and A, B, C denote matrices. The properties are conditional on the operations being defined for the case in point.

4.9.1. *Equality.*

vectors: Two n-vectors a and b are said to be equal if all their corresponding components are equal. Equality is only possible for vectors of the same dimension.

matrices: Two m x n matrices A and B are said to be equal if all their corresponding components are equal. Equality is only possible for matrices of the same dimension.

4.9.2. *Multiplication by a scalar.*

a: (α + β) A = αA + βA
b: α(A + B) = αA + αB
c: α (βA) = ($\alpha\beta$) A

Note that A and B above can be replaced by a and b as in (1)(a) = a

4.9.3. *Addition.*

a: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$
b: $\vec{a} + 0 = \vec{a}$
c: $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$
d: $\vec{a} + (-\vec{a}) = 0$
e: $A + B = B + A$
f: $A + (B + C) = (A + B) + C$
g: $A + 0 = 0 + A = A$
h: $A + (-A) = 0$

4.9.4. *Multiplication.*

a: $\vec{a}\,\vec{b} = \vec{b}\,\vec{a}$
b: AB $\neq$ BA
c: A(BC) = (AB)C
d: $\alpha(\vec{b} + \vec{c}) = \alpha\vec{b} + \alpha\vec{c}$
e: $A(B + C) = AB + AC$
f: $(B + C)A = BA + CA$
g: $(\alpha\vec{a})\vec{b} = \vec{a}(\alpha\vec{b}) = \alpha(\vec{a}\vec{b})$
h: $\vec{a} \cdot \vec{a} > 0 \Leftrightarrow \vec{a} \neq 0$
i: $\vec{a} \cdot 0 = 0 \cdot \vec{a} = 0$
j: $A0 = 0A = 0$
k: $AI = IA = A$

4.9.5. *Transposes.*

a: $(A')' = A$
b: $(ABC)' = C'\,B'\,A'$
c: $(A + B)' = A' + B'$

4.9.6. *Properties of the trace.*

a: trace (I) = n
b: trace (ABC) = trace (CAB) = trace (BCA)
c: trace (A + B) = trace (A) + trace (B)
d: tr(AB) = tr(BA) if both AB and BA are defined
e: tr(kA) = ktr(A) where k is a scalar

4.10. **Idempotent matrices.** - A matrix is called idempotent if

$$A^2 = A \tag{44}$$

For example the identity matrix is idempotent. Consider the matrix M below.

$$M = \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \tag{45}$$

We can verify that it is idempotent by carrying out the multiplication.

$$M\,M = \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \tag{46}$$

Consider the multiplication of the first row and first column

$$\begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \end{pmatrix} \begin{pmatrix} 0.8 \\ -0.2 \\ -0.2 \\ -0.2 \\ -0.2 \end{pmatrix} = 0.64 + 0.4 + 0.4 + 0.4 + 0.4 = 0.8 \tag{47}$$

Or consider the multiplication of the first row and second column

$$\begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \end{pmatrix} \begin{pmatrix} -0.2 \\ 0.8 \\ -0.2 \\ -0.2 \\ -0.2 \end{pmatrix} = -0.16 + -0.16 + 0.4 + 0.4 + 0.4 = -0.2 \tag{48}$$

Note that if A is idempotent, tr(A) = rank of A .